



Multiperiod Optimization

**Lorenz T. Biegler
Department of Chemical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213**

Bill Rooney, Carl Laird and Bart van Bloemen Waanders

**Sandia Workshop on Large-Scale Robust Optimization
September 1, 2005**



Outline

Introduction

Types of unknown information in engineering
Variability and uncertainty

Optimization under uncertainty

Monte Carlo vs. Multiperiod formulations

Extensions

Hard constraints
Control variables

Interior point multiperiod algorithm

Two stage problem formulations

Uncertainty, Variability, Both

Examples

Chemical processes
Source Inversion

Conclusions



Introduction

Goal: At design stage, incorporate changes in variable process inputs and uncertain parameters

Two types of unknown information:

What is not known well (uncertainty, here and now)...

- * Models and their parameters (kinetic and transport coefficients, etc.)
- * Unmeasured and unobservable disturbances (ambient conditions)

What is well known but is subject to change (variability, wait and see)...

- * Feed flow rates
- * Process conditions and inputs
- * Product demands
- * Changes are measured (perfectly) and control variables are used to compensate for them



Design Under Uncertainty

$$\min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(d, z, y, \theta) = 0]$$

s.t.

$$\Pr[q(d, z, y, \theta) \leq 0, d \in D, z \in Z, y \in Y, \theta \in \Theta] \geq \alpha$$

y : state variables (x , T , p , etc)

d : design variables (equipment sizes, etc)

z : control/operating variables (actuators, flows, etc)

θ : variable inputs and uncertain parameters

(no dynamics, single stage)

h : process model equations

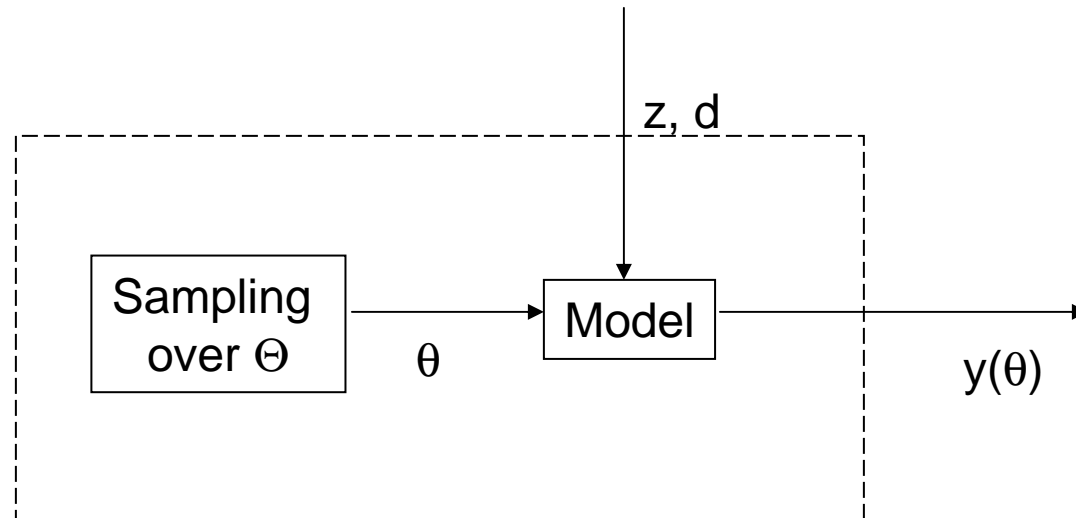
q : (some) process model inequalities

$E[P]$: expected value of an objective function

$\Pr[g]$: probability $\geq \alpha$ for chance constraints

Monte Carlo Models

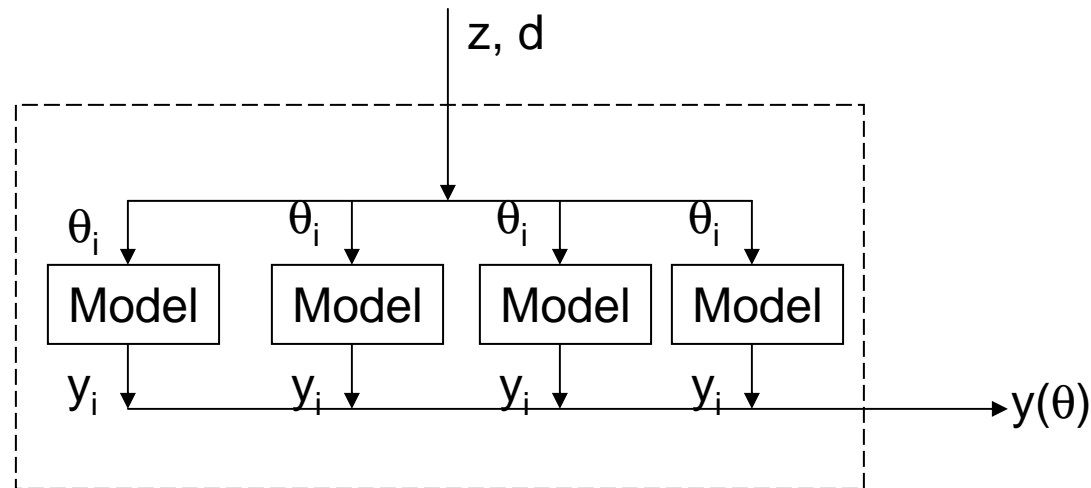
$$\begin{aligned} \min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(d, z, y, \theta) = 0] \\ \text{s.t. } Pr_{\theta}[q(d, z, y, \theta) \leq 0, d \in D, y \in Y, z \in Z, \theta \in \Theta] \geq \alpha \end{aligned}$$



Optimizer using discrete sampling over Θ
e.g., Hammersley points

Multiperiod Models

$$\begin{aligned} \min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(z, y, x, \theta) = 0] \\ \text{s.t. } Pr_{\theta}[g(d, z, y, \theta) \leq 0, d \in D, y \in Y, z \in Z, \theta \in \Theta] \geq \alpha \end{aligned}$$



Optimizer using discrete periods over Θ
e.g., Hammersley points



Multiperiod Models for Uncertainty

$$\begin{aligned} \min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(d, z, y, \theta) = 0] \\ \text{s.t. } Pr_{\theta} [q(d, z, y, \theta) \leq 0, d \in D, z \in Z, y \in Y, \theta \in \Theta] \geq \alpha \end{aligned}$$

After discretization:

$$\text{Min } f_0(d) + \sum_j \omega_j f_j(d, z, y_j, \theta_j)$$

$$\text{s.t. } h_j(d, z, y_j, \theta_j) = 0$$

$$\varphi(d, y, z) \leq 0$$

Derivation of chance constraint requires implicit quadrature formula that covers all periods, j .

What are the advantages of Multiperiod over Monte Carlo?



Multiperiod Models for Uncertainty: Addition of Hard Constraints

$$\begin{aligned} \min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(d, z, y, \theta) = 0, g(d, z, y, \theta) \leq 0] \\ \text{s.t. } Pr_{\theta} [q(d, z, y, \theta) \leq 0, d \in D, z \in Z, y \in Y, \theta \in \Theta] \geq \alpha \end{aligned}$$

After discretization:

$$\text{Min } f_0(d) + \sum_j \omega_j f_j(d, z, y_j, \theta_j)$$

$$\text{s.t. } h_j(d, z, y_j, \theta_j) = 0$$

$$g_j(d, z, y_j, \theta_j) \leq 0$$

$$\varphi(d, y, z) \leq 0$$

Hard constraints allow no violation over $\theta \in \Theta$.

Note relation to robust optimization (A. Nemirovski, Y. Zhang)

Some References: Bandoni, Romagnoli and coworkers (1993-1997), Narraway, Perkins and Barton (1991), Srinivasan, Bonvin, Visser and Palanki (2002), Walsh and Perkins (1994, 1996)



Confidence Intervals for Uncertainty

- * Uncertain model parameters often assumed to lie between lower and upper bounds and vary independently of each other

$$\theta \in \Theta = \{\theta | \theta_{low} \leq \theta \leq \theta_{up}\}$$

- * These bounds are available from confidence intervals

$$\Theta = \{\theta | \theta = \hat{\theta} \pm \sigma t_{1-\frac{\alpha}{2}, n-p}\}$$

- σ : Standard deviation of each parameter
- t : Student's t distribution
- α : Confidence level
- p : Number of uncertain parameters θ
- n : Number of data points



Ellipsoidal Confidence Regions for Uncertainty

Replace hypercube with elliptical confidence regions:

$$\Theta = \{\theta | (\theta - \hat{\theta})^T V_{\theta}^{-1} (\theta - \hat{\theta}) \leq p F_{(1-\alpha, p, n-p)}\}$$

V_{θ} : Covariance matrix

F : Value of the F distribution

- * p -dimensional ellipsoidal region
- * Attempts to cover all *joint* parameter combinations
- * Quadratic (convex) constraint in θ
- * Approximate for nonlinear systems

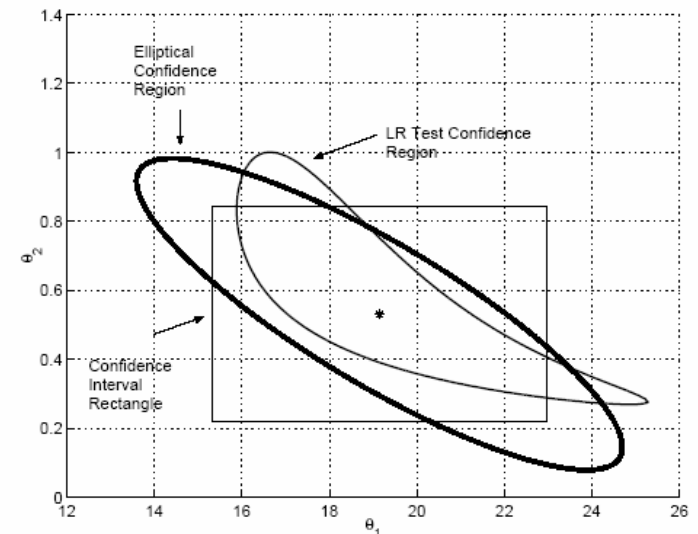
Question: How to describe confidence regions for *nonlinear* problems?

Nonlinear Confidence Regions for Uncertainty

Replace ellipse with confidence regions from the Likelihood Ratio Test:

$$\Theta = \{\theta | 2[L(\theta^*) - L(\theta)] \leq \eta \chi_{p,1-\alpha}^2\}$$

- θ^* : Maximum likelihood estimates
- L : Log-likelihood function
- η : Bartlett correction factor accounts for finite experimental data size
- χ^2 : Chi-squared statistic



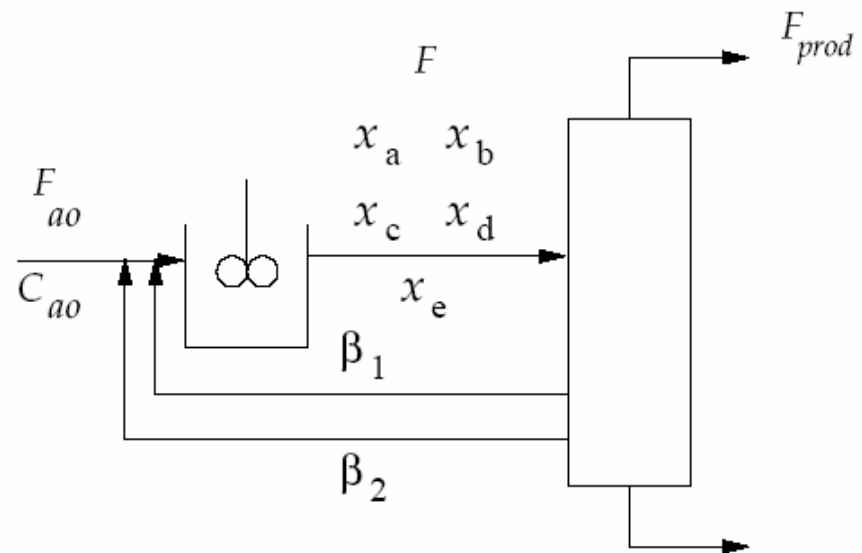
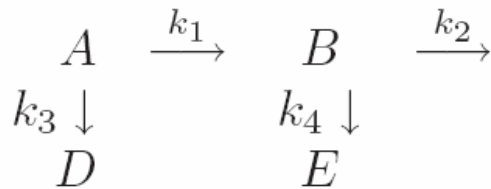
$$y = \theta_1(1 - \exp(\theta_2 t))$$

- * Contours of $L(\theta)$ map out the confidence region
- * Response functions and the data help form the confidence regions



Process Example (here and now)

- * Problem: Minimize the cost (V, F) to produce desired product
- * Denbigh's reaction takes place with uncertainty in each rate constant

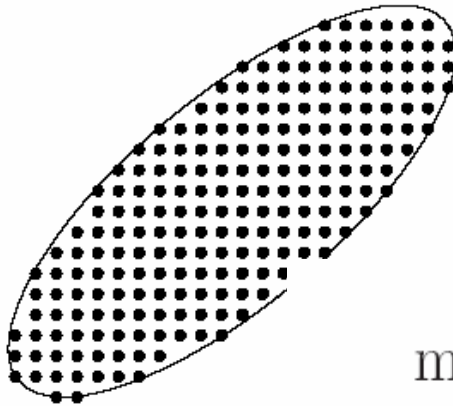


Reactor-Separator Flowsheet



Process Example (here and now)

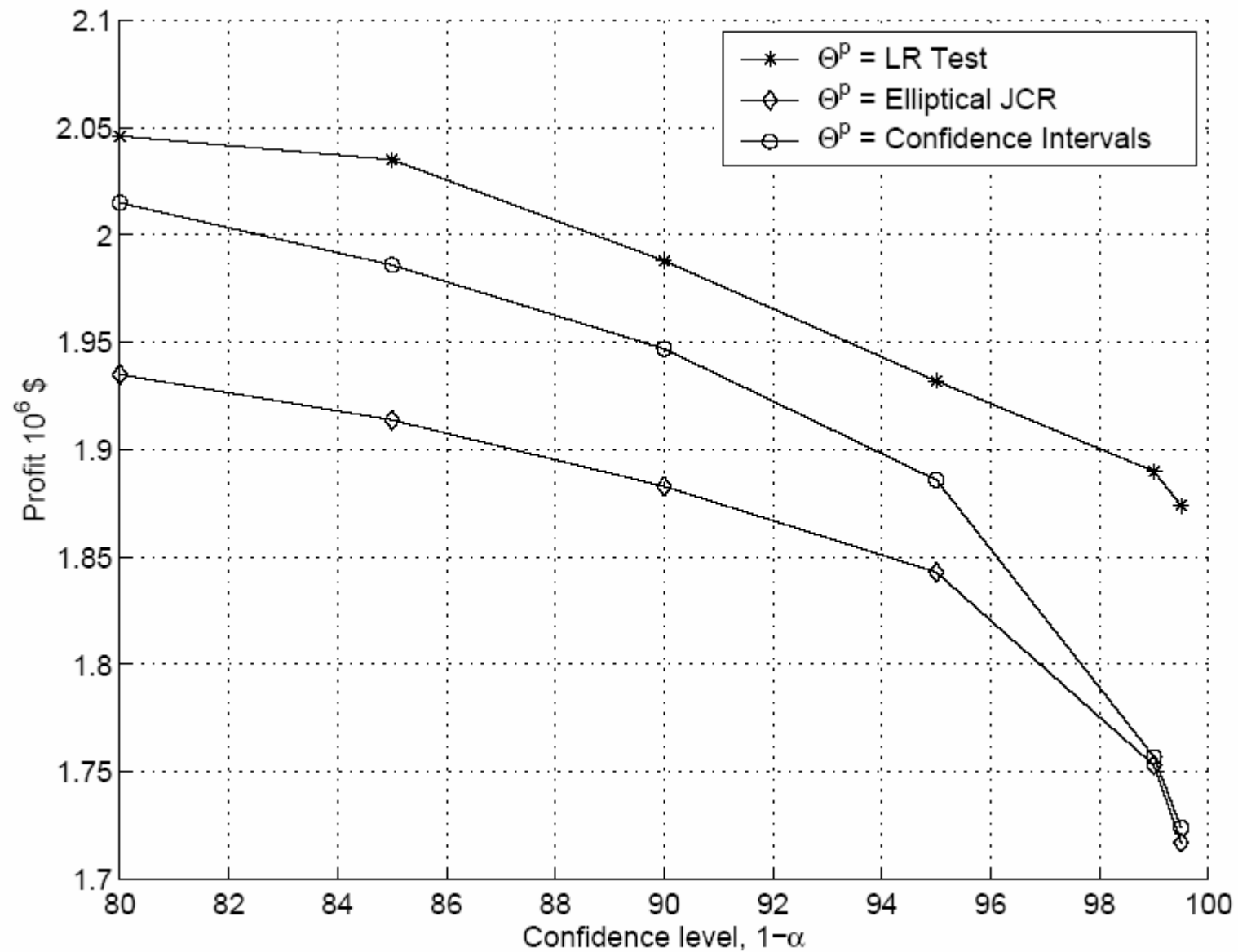
- * Optimize an estimated average profit over each confidence region
- * 2400 points sampled from each confidence region



$$\begin{aligned} \max \quad & \frac{1}{2400} \sum_{i=1}^{2400} P_i(d, u, x_{c_i}, \theta_i) \\ \text{s.t.} \quad & h(d, u, x_i, \theta_i) = 0 \\ & g(d, u, x_i, \theta_i) \leq 0 \\ & \theta_i \in \Theta^p; \quad i = 1, \dots, 2400 \end{aligned}$$

- * How does profit change with confidence level, α

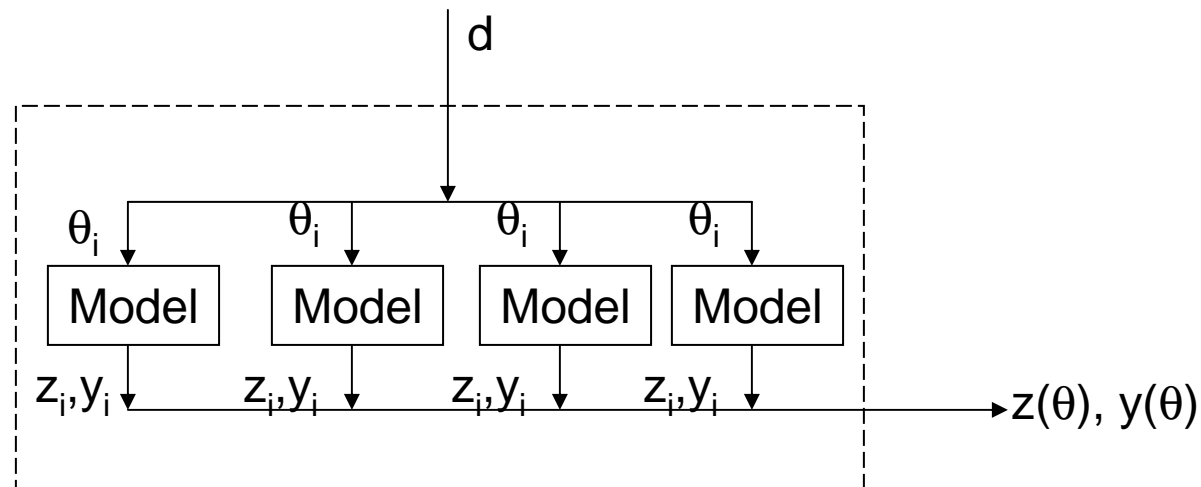
Influence of distributions on profit





Addressing Variability (wait and see)

- * Process parameters (temperatures, pressures, etc.) with known changes during plant operation
- * Changes are measured (perfectly) and control variables are used to compensate for them (recourse)
- * Control variables are used to improve the results in the design problems, can be adjusted as soon as variability is known
- * Parameters, $\theta_v \in \Theta_v$, account for process variability, not uncertainty





Multiperiod Models for Variability

$$\begin{aligned} \min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(d, z, y, \theta) = 0, g(d, z, y, \theta) \leq 0] \\ \text{s.t. } Pr_{\theta} [q(d, z, y, \theta) \leq 0, d \in D, z \in Z, y \in Y, \theta \in \Theta] \geq \alpha \end{aligned}$$

After discretization:

$$\text{Min } f_0(d) + \sum_j \omega_j f_j(d, z_j, y_j, \theta_j)$$

$$\text{s.t. } h_j(d, z_j, y_j, \theta_j) = 0$$

$$g_j(d, z_j, y_j, \theta_j) \leq 0$$

$$\varphi(d, y, z) \leq 0$$

Control variables offer more freedom to deal with variability
(e.g., reject disturbances)

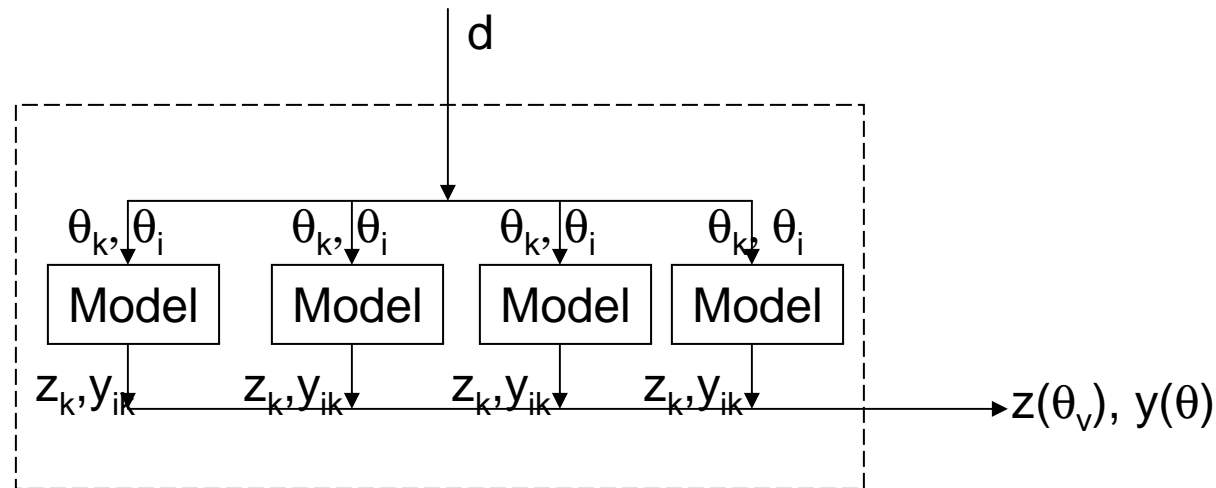
Some References: Grossmann and coworkers (1983-1991), Ierapetritou, Acevedo and Pistikopoulos (1996), Pistikopoulos and coworkers (1995-2001)



Incorporating both uncertainty and variability

Control variables:

- * Allowed to compensate for varying process parameters θ_v (e.g., measured disturbances)
- * Not allowed to compensate for uncertainty model parameters, θ_p (kinetic and transport parameters)
- * z_k indexed by θ_v but not by θ_p
- * y_{ik} indexed by θ_v and by θ_p





Multiperiod Models for Both

$$\begin{aligned} \min E_{\theta}[P(d, z, y, \theta) \text{ s.t. } h(d, z, y, \theta) = 0, g(d, z, y, \theta) \leq 0] \\ \text{s.t. } Pr_{\theta} [q(d, z, y, \theta) \leq 0, d \in D, z \in Z, y \in Y, \theta \in \Theta] \geq \alpha \end{aligned}$$

After discretization:

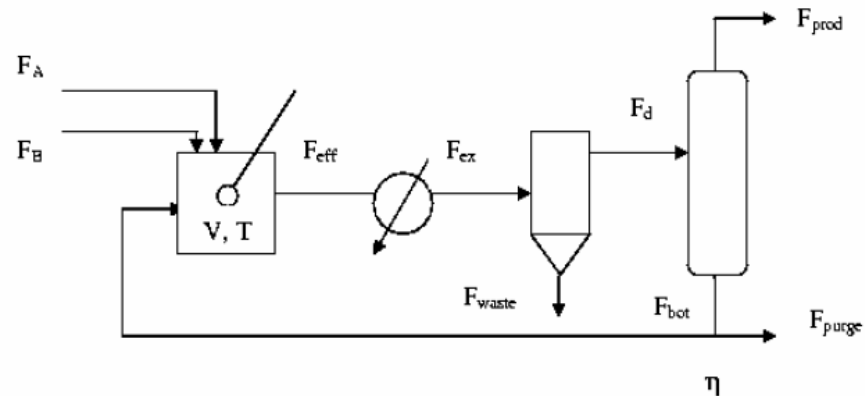
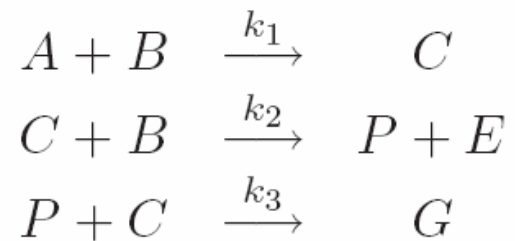
$$\begin{aligned} \text{Min } f_0(d) + \sum_{i,k} \omega_{ik} f_{ik}(d, z_k, y_{ik}, \theta_{v,k}, \theta_{p,i}) \\ \text{s.t. } h_{ik}(d, z_k, y_{ik}, \theta_{v,k}, \theta_{p,i}) = 0 \\ g_{ik}(d, z_k, y_{ik}, \theta_{v,k}, \theta_{p,i}) = 0 \\ \varphi(d, y, z) \leq 0 \end{aligned}$$

Control variables offer freedom to deal with variability
(e.g., reject disturbances) but not uncertainty



Uncertainty and Variability: Williams-Otto Process (Rooney, B., 2003)

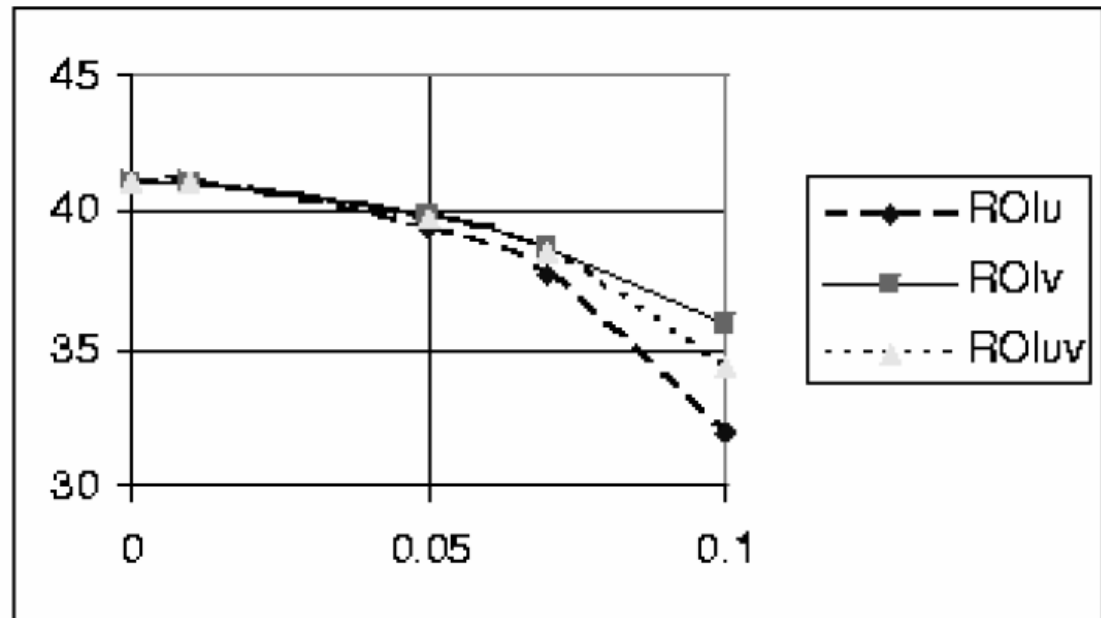
- Problem: Maximize ROI to produce product P
- Series reactions with rate constants uncertain



- Uncertain model parameters, a_1, a_2 and a_3
- Varying process parameters: $F_A = 10000(1 \pm \delta)$ and $F_B = 40000(1 \pm \delta)$

Williams-Otto Results

Conf. Reg.	δ	Θ^u
Nominal	0	41.15
Hypercube	0.01	41.08
	0.05	39.42
	0.07	37.80
	0.10	31.92
Ellipse	0.01	41.092
	0.05	39.621
	0.07	38.176
	0.10	34.481



- Treating uncertainty and variability separately gives intermediate and 'more realistic' results
- Using elliptical confidence regions for θ^p has a strong influence on cost



Interior Point Method

$$\text{Min } f_0(d) + \sum_j \omega_j f_j(d, z_j, y_j, \theta_j)$$

$$\text{s.t. } h_j(d, z_j, y_j, \theta_j) = 0$$

$$g_j(d, z_j, y_j, \theta_j) + s_j = 0$$

$$\varphi(d, y, z) + \sigma = 0, \quad \sigma, s_j \geq 0$$

$$\text{Min } f_0(p) + \sum_j \omega_j f_j(p, x_j)$$

$$\text{s.t. } c_j(p, x_j) = 0$$

$$\bar{c}(p, x) = 0, \quad p, x_j \geq 0$$

$$\text{Min } f_0(p) + \sum_j \omega_j f_j(p, x_j) - \mu \left\{ \sum_{j,l} \ln x_j^l + \sum_{j,l} \ln p^l \right\}$$

$$\text{s.t. } c_j(p, x_j) = 0$$

$$\bar{c}(p, x) = 0$$

$$\mu^i \rightarrow 0 \Rightarrow [x(\mu^i), p(\mu^i)] \rightarrow [x^*, p^*]$$



IPOPT Algorithm

• *Line Search Strategies*

- - ℓ_2 exact penalty merit function
- - augmented Lagrangian merit function
- - **Filter method (adapted and extended from Fletcher and Leyffer)**

• *Hessian Calculation*

- - BFGS (full/LM and reduced space)
- - SR1 (full/LM and reduced space)
- - **Exact full Hessian (direct)**
- - Exact reduced Hessian (direct)
- - Preconditioned CG

• *Freely Available*

- CPL License and COIN-OR distribution
- Solved on 1000s of test problems and applications
- Recently rewritten in C++
- Code available at <http://www.coin-or.org>

• *Algorithmic Properties*

- Globally and superlinearly convergent (see Wächter, B., 2005)
- Weaker assumptions than other codes
- ***Easily tailored to different problem structures***



Optimality Conditions: Interior point formulation

Define Lagrange Function:

$$L(x, p) = f_0(p) + \bar{c}(x, p)^T \bar{\lambda} - \underline{v_p}^T p + \sum_i [\omega_i f_i(x_i, p) + c_i(x_i, p)^T \lambda_i - \underline{v_i}^T x_i]$$

Take Stationary Conditions:

$$\begin{aligned} \omega_i \nabla_{x_i} f_i(x_i, p) + \nabla_{x_i} c_i(x_i, p) \lambda_i + \nabla_{x_i} \bar{c}(x, p) \bar{\lambda} - v_i &= 0 \\ \nabla_p f_0(p) + \sum_i [\omega_i \nabla_p f_i(x_i, p) + \nabla_p c_i(x_i, p) \lambda_i] + \nabla_p \bar{c}(x, p) \bar{\lambda} - v_p &= 0 \\ X_i V_i e - \mu e &= 0 \\ P V_p e - \mu e &= 0 \\ c(x_i, p) &= 0 \\ \bar{c}(x, p) &= 0 \end{aligned}$$



Newton Step for IPOPT

$$\begin{bmatrix} W_1 & & & & & \\ & W_2 & & & & \\ & & W_3 & & & \\ & & & \dots & & \\ & & & & W_N & \\ w_1^T & w_2^T & w_3^T & \dots & w_N^T & W_p \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_N \\ u_p \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \\ r_p \end{bmatrix}$$

$$W_i = \begin{bmatrix} (\nabla_{x_i, x_i} L^k + (X_i^k)^{-1} V_i^k) & \nabla_{x_i} c_i(x_i^k, p^k) \\ \nabla_{x_i} c_i(x_i^k, p^k)^T & 0 \end{bmatrix} \quad u_i = \begin{bmatrix} \Delta x_i \\ \Delta \lambda_i \end{bmatrix} \quad u_p = \begin{bmatrix} \Delta p \\ \Delta \bar{\lambda} \end{bmatrix}$$

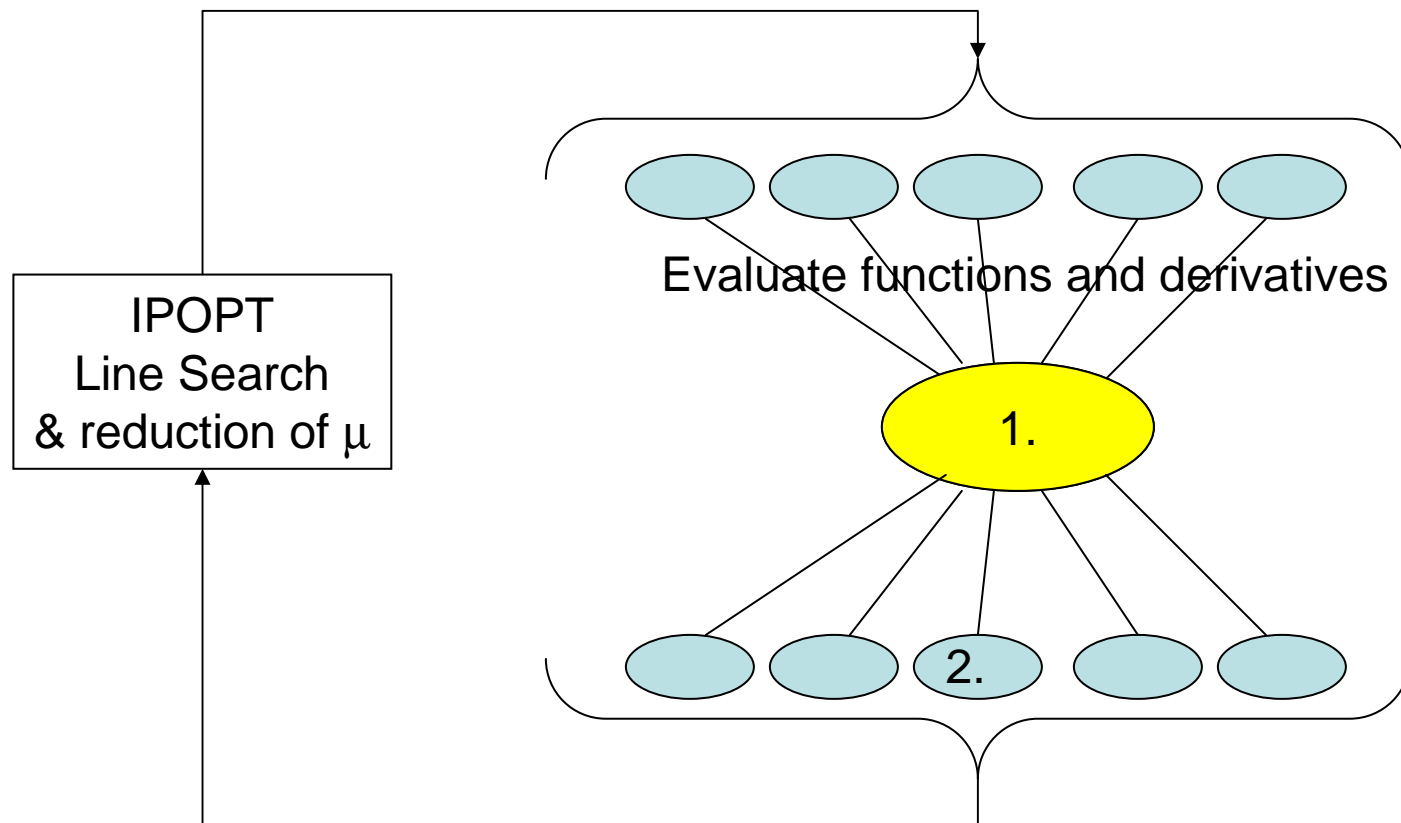
$$W_p = \begin{bmatrix} \nabla_{p,p} L^k + (P^k)^{-1} V_p^k & \nabla_p \bar{c} \\ \nabla_p \bar{c}^T & 0 \end{bmatrix} \quad w_i = \begin{bmatrix} W_{x_i p} & \nabla_{x_i} \bar{c} \\ \nabla_p c_i^T & \end{bmatrix}$$

Decomposition Algorithm

Key Steps

$$1. \left(W_{pp} - \sum_i w_i^T W_i w_i \right) \Delta u_p = r_p - \sum_i w_i^T W_i r_i$$

$$2. W_i \Delta u_i = r_i - w_i \Delta u_p$$

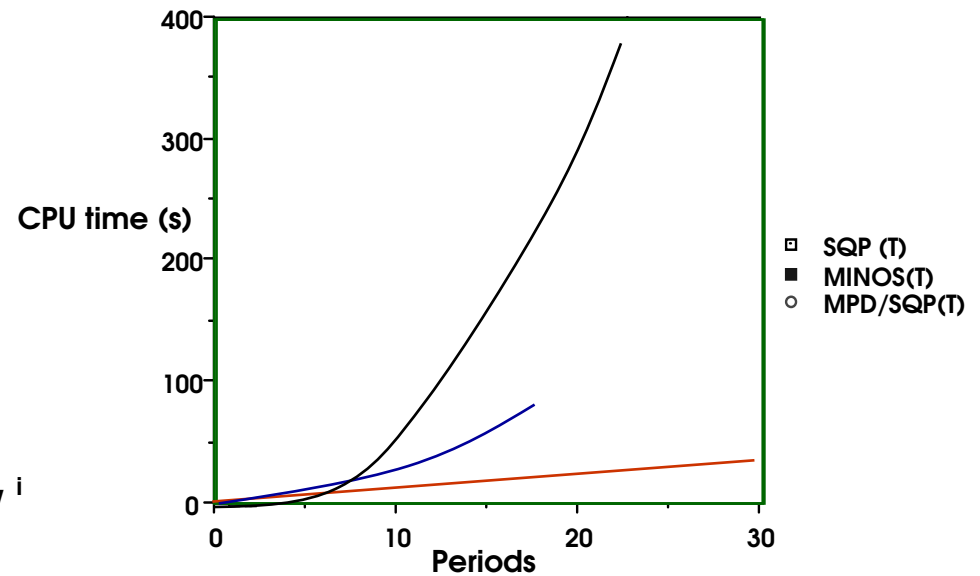
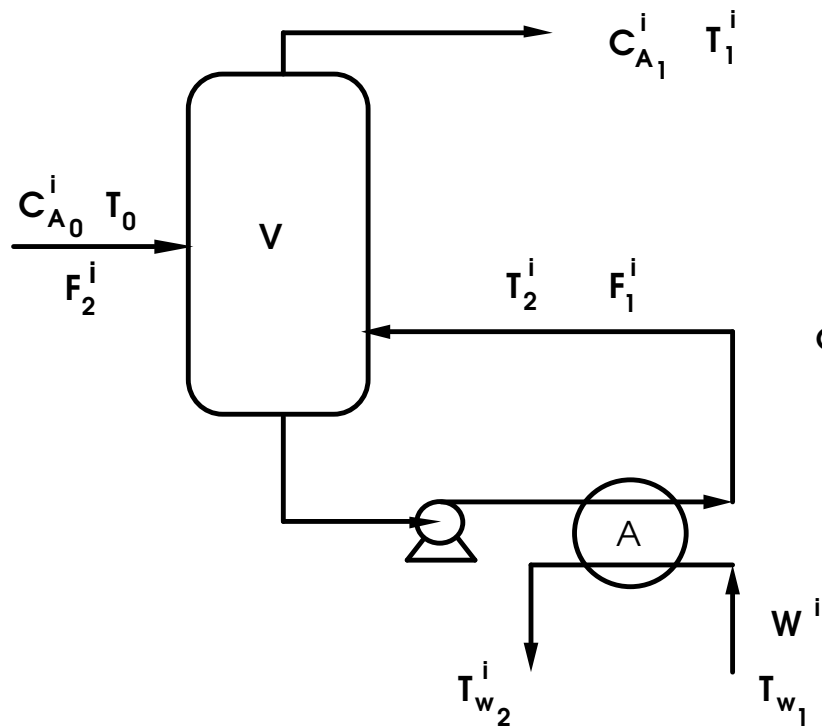


Computational cost is linear in number of periods
Trivial to parallelize

Multiperiod Flowsheet 1

(13+2) variables and (31+4) constraints (1 period)

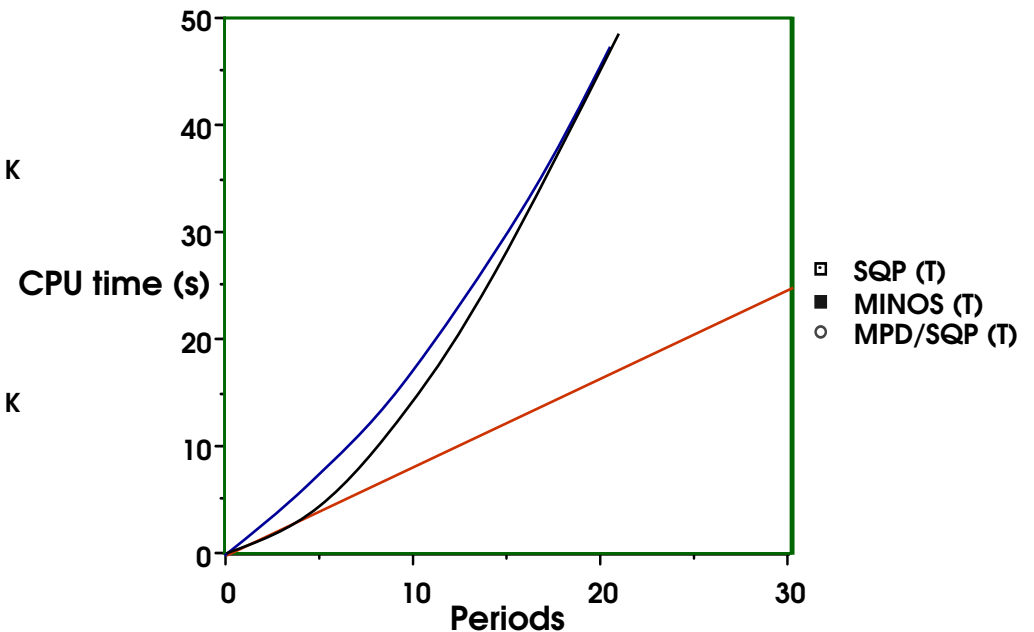
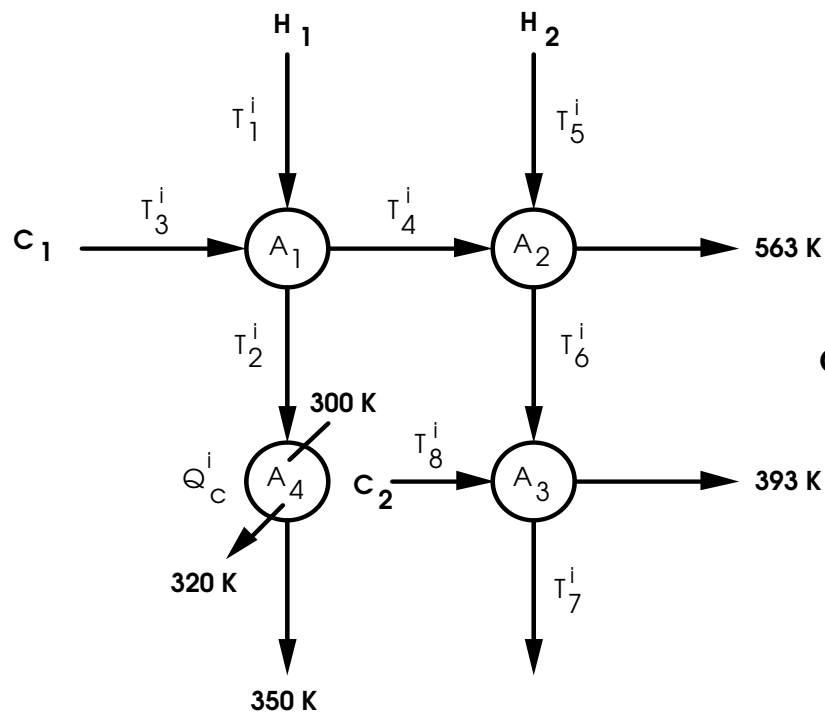
262 variables and 624 constraints (20 periods)



Multiperiod Example 2 – Heat Exchanger Network

(12+3) variables and (31+6) constraints (1 period)

243 variables and 626 constraints (20 periods)

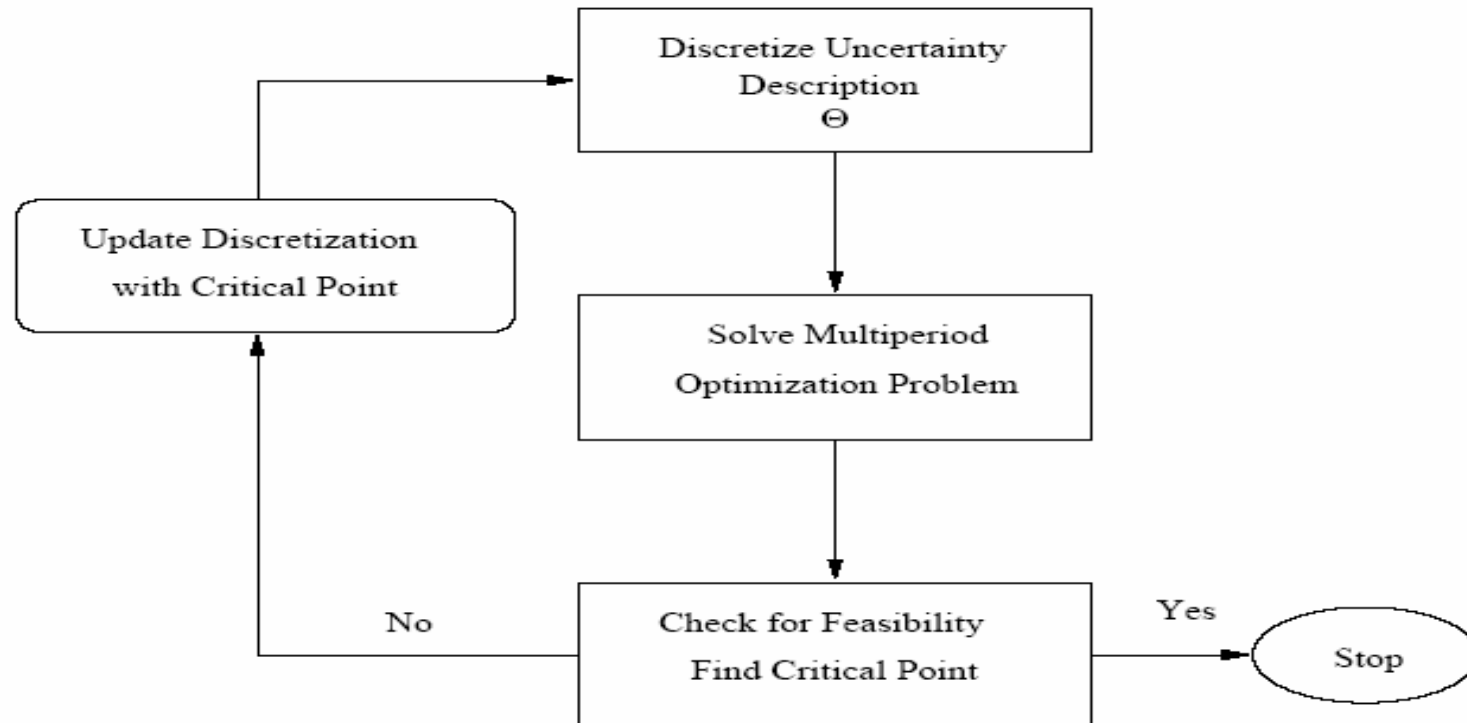




Hard Constraints

- Need to enforce hard constraints over entire domain, $\theta \in \Theta$.
- Sampling distribution to approximate E and Pr operators is not a guarantee.
- Define $T(d) \leq 0$ to represent feasibility over $\theta \in \Theta$
- What does this function look like?
- How do we incorporate this into the algorithm?

“Two Stage” Algorithm



- Solve multiperiod problem for $\theta_j \in \Theta$ to yield a given design
- Attempt to 'break' the design with a feasibility test \rightarrow locates a critical θ
- Add critical θ to multiperiod problem and solve again



Feasibility Tests

Feasibility problem for parameter uncertainty (Conservative):

$$\forall \theta \in \Theta \forall j \{g_j(d, z, \theta) \leq 0\} \Rightarrow \text{Max}_{\theta \in \Theta} \text{Max}_j \{g_j(d, z, \theta) \leq 0\}$$

Feasibility problem for variability (Optimistic):

$$\forall \theta \in \Theta \exists z \in Z \forall j \{g_j(d, z, \theta) \leq 0\} \Rightarrow \text{Max}_{\theta \in \Theta} \text{Min}_{z \in Z} \text{Max}_j \{g_j(d, z, \theta) \leq 0\}$$

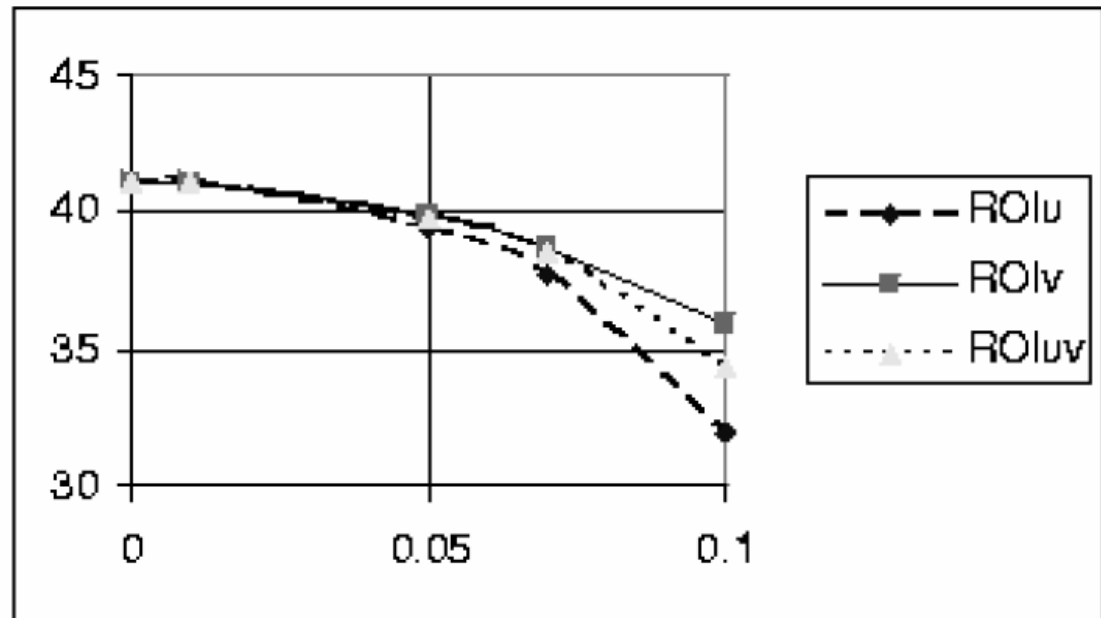
Feasibility problem for variability and uncertainty (Realistic):

$$\forall \theta_v \in \Theta_v \exists z \in Z \forall \theta_p \in \Theta_p \forall j \{g_j(d, z, \theta) \leq 0\} \Rightarrow \\ \text{Max}_{\theta_v \in \Theta_v} \text{Min}_{z \in Z} \text{Max}_{\theta_p \in \Theta_p} \text{Max}_j \{g_j(d, z, \theta) \leq 0\}$$

- Global solutions required for each operator (see Swaney and Grossmann (1985) for properties and analysis)
- Nested problems solved by writing optimality conditions at multiple levels – leads to difficulties for NLPs (specific convexity properties required).
 - KS function aggregation (Rooney and Biegler, 2003)
 - Branch and Bound Search (Achenie and Ostrovsky, 2003)
 - Global algorithms (Ierapetritou, Floudas...)

Williams-Otto Results: satisfies all feasibility tests

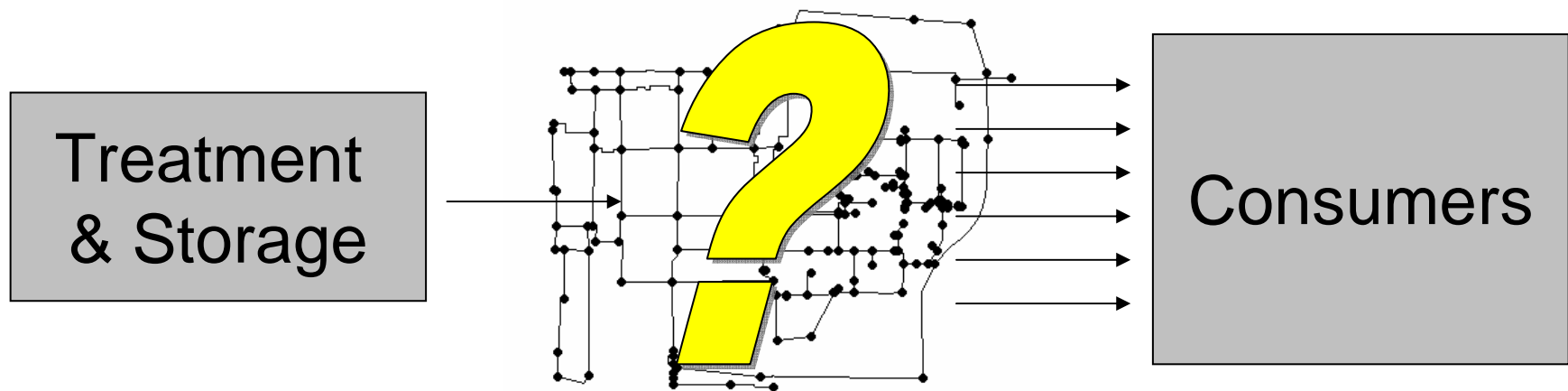
Conf. Reg.	δ	Θ^u
Nominal	0	41.15
Hypercube	0.01	41.08
	0.05	39.42
	0.07	37.80
	0.10	31.92
Ellipse	0.01	41.09
	0.05	39.621
	0.07	38.176
	0.10	34.481



- Treating uncertainty and variability separately gives intermediate and 'more realistic' results
- Using elliptical confidence regions for θ^p has a strong influence on cost



Future Work: Source Detection in Municipal Water Networks with Uncertainty



- Large Area Encompassed
 - Many, Many Access Points
- }
- Vulnerable to
 - Accidental & Intentional
 - Contamination



Optimization Problem

Node Concentrations &
Injection Terms Only

$$\min_{m(t), \bar{c}(x,t), \hat{c}(t)} \psi = \sum_{r \in \Theta_s} \sum_{k \in \mathcal{N}_s} \frac{1}{2} \int_0^{t_f} w_k(t) (\hat{c}_k(t) - \hat{c}_k^*(t))^2 \delta(t-t_r) dt + \frac{\rho}{2} \int_0^{t_f} m_k(t)^2 dt$$

$$\left. \begin{aligned} \frac{\partial \bar{c}_i(x,t)}{\partial t} + u_i(t) \frac{\partial \bar{c}_i(x,t)}{\partial x} &= 0, \\ \bar{c}_i(x=\mathcal{I}_i(t), t) &= \hat{c}_{k_i(t)}(t), \\ \bar{c}_i(x, t=0) &= 0, \end{aligned} \right\} \forall i \in \mathcal{P},$$

Only Constraints
with Spatial
Dependence

$$\hat{c}_k(t) = \frac{\left(\sum_{i \in \Gamma_k(t)} Q_i(t) \bar{c}_i(x=\mathcal{O}_i(t), t) \right) + m_k(t)}{\left(\sum_{i \in \Gamma_k(t)} Q_i(t) \right) + Q_k^{ext}(t) + Q_k^{inj}(t)}, \quad \forall k \in \mathcal{J},$$

Pipe Boundary
Concentrations

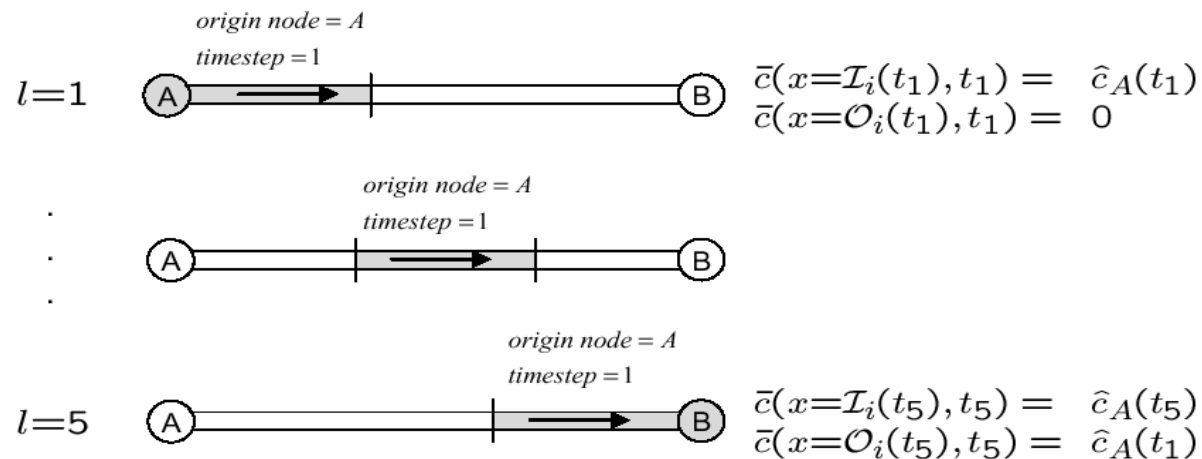
$$\left. \begin{aligned} V_k(t) \frac{d\hat{c}_k(t)}{dt} &= \left(\sum_{i \in \Gamma_k(t)} Q_i(t) \bar{c}_i(x=\mathcal{O}_i(t), t) \right) + m_k(t) - \left[\left(\sum_{i \in \Gamma_k(t)} Q_i(t) \right) + Q_k^{ext}(t) + Q_k^{inj}(t) \right] \hat{c}_k(t), \\ \hat{c}_k(t=0) &= 0, \end{aligned} \right\} \forall k \in \mathcal{S},$$

$$m_k(t) \geq 0, \quad \forall k \in \mathcal{N}.$$

Injection Terms Only



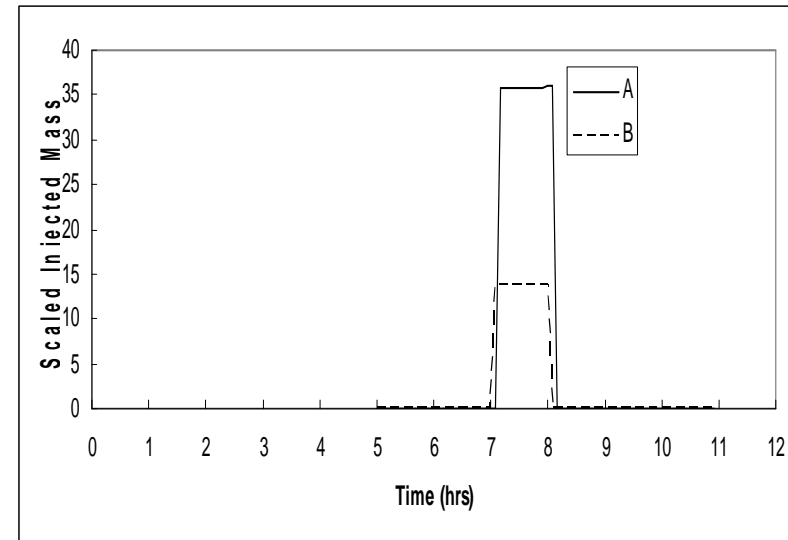
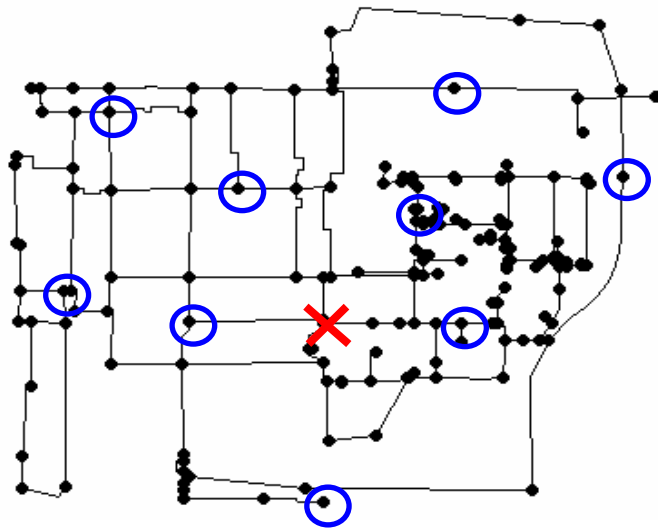
Origin Tracking Algorithm



- Known Hydraulics – Function of Time
- Pipe Network PDEs Linear in Concentration
- Pipe by Pipe PDEs
 - Efficient for Large Networks
 - Convert PDEs to DAEs with variable time delays
- Removes Need to Discretize in Space
- Discretization in time leads to a large QP

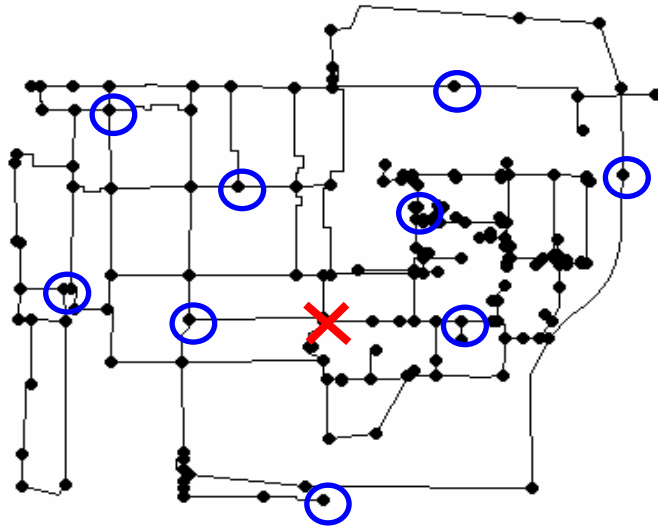


Municipal Source Detection Example

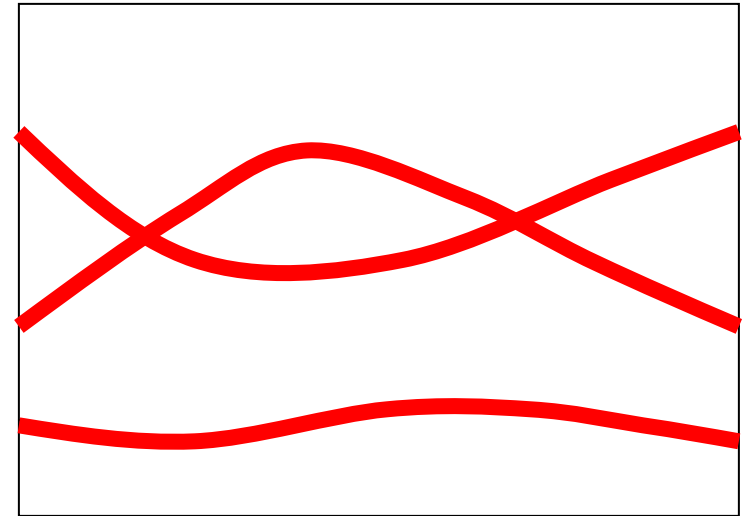


- Algorithm successful on over 1000 numerical tests with real municipal water networks
- Solution time < 2 CPU minutes for ~ 250,000 variables, ~45,000 degrees of freedom
 - Effective in a real time setting
- Formulation tool links to existing water network software
- Can impose unique solutions through an extended MIQP formulation (post-processing phase)

Source Detection with Uncertainty



Diurnal network hydraulics



- Incorporate uncertain demands in diurnal hydraulics (somewhat simplistic)
- Find injection location $m_k(t)$ as “design variable”
- Formulate as multiperiod problem and apply algorithm
- Exploit properties from IPOPT and BBD decomposition
- Impose unique solutions through an extended MIQP formulation (same as in single scenario, same effort)
- Hard constraints ($c_k(t) \geq 0$) but $m_k(t) \geq 0 \Rightarrow c_k(t) \geq 0$
- No feasibility problems needed
- Other applications: robust sensor placement solved with larger MINLPs?



Conclusions

- Combined Variability and Uncertainty
 - Overdesign for uncertainty, θ_p
 - Apply (feedforward) control for variability θ_v
- Multiperiod Problem
 - Scenarios for θ_p and/or θ_v considered simultaneously
 - BBD structure exploited by IPOPT algorithm
 - computational cost linear in scenarios (nearly perfect speedup if parallelized)
- Modified two-stage formulation
 - Control variables indexed only for variability in multiperiod problem
 - More challenging feasibility problems
 - Yields intermediate results: less conservative and not overly optimistic
 - Identification of θ_v and θ_p and corresponding distributions is still an open question.